Dans tout le problème, n désigne un entier naturel non nul.
On considère une urne U_n contenant n boules numérotées de 1 à n. On tire une boule au hasard dans U_n . On note k le numéro de cette boule. Si k est égal à 1, on arrête les tirages. Si k est supérieur ou égal à 2, on enlève de l'urne U_n les boules numérotées de k à n (il reste donc les boules numérotées de 1 à $k-1$), et on effectue à nouveau un tirage dans l'urne. On répète ces tirages jusqu'à l'obtention de la boule numéro 1. On note X_n la variable aléatoire égale au nombre de tirages nécessaires pour l'obtention de la boule numéro 1. On note Y_n la variable aléatoire égale à la somme des numéros des boules tirées. On note $E(X_n)$ et $V(X_n)$ (respectivement $E(Y_n)$ et $V(Y_n)$) l'espérance et la variance de X_n (respectivement Y_n).

Partie 1.

1. On pose : $h_n = \sum\limits_{k=1}^{n} \dfrac{1}{k} = 1 + \dfrac{1}{2} + ... + \dfrac{1}{n}$

a) On peut utiliser l'inégalité des accroissements finis :
$\ln'(x) = \frac{1}{x}$ et pour $x \in [k, k+1]$ on a $\frac{1}{k+1} \leq \ln'(x) \leq \frac{1}{k}$ donc comme $k \leq k+1$ on a

$$\frac{1}{k+1} = \frac{1}{k+1}(k+1-k) \leqslant \ln(k+1) - \ln k \leqslant (k+1-k)\frac{1}{k} = \frac{1}{k}$$

ou de façon plus élémentaire prouver séparément les deux inégalités en étudiant les variations de la différence.

b) On a donc en sommant les inégalité :

$$\sum_{k=1}^{n} \frac{1}{k} \geq \sum_{k=1}^{n} (\ln(k+1) - \ln k)$$

somme qui se simplifie en cascade : $\sum\limits_{k=1}^{n} (\ln(k+1) - \ln k) = \sum\limits_{k=1}^{n} \ln(k+1) - \sum\limits_{k=1}^{n} \ln k$ et en réindexant
$= \sum\limits_{k=2}^{n+1} \ln k - \sum\limits_{k=1}^{n} \ln k = \ln(n+1) - \ln(1)$ d'où finalement :

$$\sum_{k=1}^{n} \frac{1}{k} \geq \ln(k+1)$$

On réutilise $\dfrac{1}{k+1} \leqslant \ln(k+1) - \ln k$ pour tout $k \geq 1$ en substituant $k-1$ à k

On a donc pour tout $k \geq 2$: $\dfrac{1}{k} \leq \ln(k) - \ln(k-1)$ que l'on ne peut utiliser que pour $k \geq 2$

$$\begin{aligned}\sum_{k=1}^{n} \frac{1}{k} &= 1 + \sum_{k=2}^{n} \frac{1}{k} \leq 1 + \sum_{k=2}^{n} (\ln(k) - \ln(k-1)) \\ &\leq 1 + \ln(n)\end{aligned}$$

D'où le résultat recherché : $\ln(n+1) \leqslant h_n \leqslant 1 + \ln n$

c) En factorisant par le prépondérant on trouve :

$$\ln(n+1) = \ln(n(1+1/n)) = \ln(n) + \ln(1+1/n)$$

et en divisant par $\ln(n)$:

$$\begin{aligned} \ln(n) + \ln(1+1/n) &\leqslant h_n \leqslant 1 + \ln n \\ 1 + \frac{\ln(1+1/n)}{\ln(n)} &\leqslant \frac{h_n}{\ln(n)} \leqslant 1 + \frac{1}{\ln(n)} \end{aligned}$$

donc par encadrement $\dfrac{h_n}{\ln(n)} \to 1$ et $h_n \sim \ln n$ quand n tend vers l'infini.

2. On pose : $k_n = \sum\limits_{k=1}^{n} \dfrac{1}{k^2} = 1 + \dfrac{1}{2^2} + ... + \dfrac{1}{n^2}$

a) Pour tout $k \geq 2$, on calcule la différence que l'on factorise :

$$\frac{1}{k^2} - \frac{1}{k-1} + \frac{1}{k} = -\frac{1}{k^2(k-1)} \leq 0$$

donc $\dfrac{1}{k^2} \leq \dfrac{1}{k-1} - \dfrac{1}{k}$

b) On somme les inégalités pour $k \geq 2$:

$$\begin{aligned} \sum_{k=1}^{n} \frac{1}{k^2} &= 1 + \sum_{k=2}^{n} \frac{1}{k^2} \leq 1 + \sum_{k=2}^{n} \left(\frac{1}{k-1} - \frac{1}{k} \right) \\ &\leq 1 + \frac{-1}{n} + \frac{1}{2-1} \leq 2 \end{aligned}$$

(la somme se simplifie en cascade comme précédemment, on a détaillé le calcul une fois, il est inutile de le refaire chaque fois)

Donc $k_n \leqslant 2$

c) En factorisant par le prépondérant, $\ln(n)$, on trouve $h_n - k_n = \ln(n) \left(\dfrac{h_n}{\ln(n)} - \dfrac{k_n}{\ln(n)} \right)$

Et comme $0 \leq k_n \leqslant 2$ on a alors $0 \leq \dfrac{k_n}{\ln(n)} \leq \dfrac{2}{\ln(n)}$ et par encadrement $\frac{k_n}{\ln(n)} \to 0$ donc la parenthèse tende vers 1 et $h_n - k_n \sim \ln(n)$ quand n tend vers l'infini.

Partie 2 : Etude de la variable aléatoire X_n

On note I_n la variable aléatoire égale au numéro de la première boule tirée dans l'urne U_n.

1. **a)** I_n est le numéro de la première boule tirée. Comme elles sont toutes équiprobables, I_n suit une loi uniforme sur $[[1, n]]$. Pour tout $k \in [[1, n]]$, $\mathrm{P}(I_n = k) = 1/n$

b) Quand $I_n = 1$, on obtient la boule 1 au premier tirage, donc $X_n = 1$; et la loi conditionnelle de X_n est : $\mathrm{P}(X_n = 1 / I_n = 1) = 1$

c) Quand $I_n = k \geq 2$, on obtient k au premier tirage et on retire toutes les boules de numéros supérieur à k.

On continue donc à partir du 2° tirage avec $k - 1 \geq 1$ boules.

Pour obtenir 1 au $j^{ème}$ tirage, il reste donc $j - 1$ tirages à effectuer (à partir du 2°) avec $k - 1$ boules.

Donc

$$\forall j \in \mathbb{N}^*,\ \forall k \in \{2\ , ..., n\}\,,\quad \mathrm{P}_{I_n=k}\left(X_n = j\right) = \mathrm{P}\left(X_{k-1} = j - 1\right)$$

2. a) Pour la loi de X_1, on n'a qu'une seule boule : la numéro 1.

Donc on l'obtient dès le premier tirage, et la loi deX_1 est : $X_1(\Omega) = \{1\}$ et $\mathrm{P}(X_1 = 1) = 1$

b) Pour X_2 on dispose au départ des boules 1 et 2. $(X_2 = 1)$ est l'événement "obtenir 1 au premier tirage".

Donc $(X_2 = 1) = (I_2 = 1)$ et $\mathrm{P}(X_2 = 1) = 1/2$

Si on n'a pas 1 au premier tirage, on aura 2 et il ne restera que la 1 dans l'urne. On obtiendra alors 1 au second tirage.

Donc $\mathrm{P}(X_2 = 2) = \mathrm{P}(I_2 = 2) = 1/2$ et finalement :

k	1	2	
$\mathrm{P}(X_2 = k)$	1/2	1/2	
$k\,\mathrm{P}(X_2 = k)$	1/2	1	$3/2 = E(X_2)$
$k^2\,\mathrm{P}(X_2 = k)$	1/2	2	$5/2 = E(X_2^2)$

d'où $V(X_2) = \dfrac{5}{2} - \dfrac{9}{4} = \dfrac{1}{4}$

c) Quand $I_3 = 1$, on a la boule 1 dès le premier tirage et donc $X_3 = 1$. Donc $\mathrm{P}(X_3 = 2/I_3 = 1) = 0$

Quand $I_3 = 2$ on a la boule 2 au premier tirage, donc pour le second il ne reste dans l'urne que la boule 1; on est donc sûr de l'obtenir au second. Donc $\mathrm{P}(X_3 = 2/I_3 = 2) = 1$.

Quand $I_3 = 3$, il reste les boules 1 et 2 dans l'urne pour le second tirage. La probabilité d'obtenir 1 au second tirage est donc 1/2 et $\mathrm{P}(X_3 = 2/I_3 = 3) = 1/2$

On utilise alors la formule des probabilités totales : $I_3 = 1$, 2 ou 3 est un système complet d'événements donc

$$\begin{aligned}
\mathrm{P}(X_3 = 2) &= \mathrm{P}_{I_3=1}(X_3 = 2)\,\mathrm{P}(I_3 = 1) + \mathrm{P}_{I_3=2}(X_3 = 2)\,\mathrm{P}(I_3 = 2) \\
&\quad + \mathrm{P}_{I_3=3}(X_3 = 2)\,\mathrm{P}(I_3 = 3) \\
&= 0 + 1\frac{1}{3} + \frac{1}{2}\frac{1}{3} = \frac{1}{2}
\end{aligned}$$

On sait déjà que $\mathrm{P}(X_3 = 1) = 1/3$ donc (loi d'une variable aléatoire) $\mathrm{P}(X_3 = 3) = 1 - \mathrm{P}(X_3 = 1) - \mathrm{P}(X_3 = 2) = 1/6$

(on pouvait aussi décomposer $X_3 = 3$: on doit avoir 1 au $3^{ème}$ et 3 au premier sinon on ne fait que 2 tirages donc 2 au 2°)

k	1	2	3	
$\mathrm{P}(X_3 = k)$	1/3	1/2	1/6	1
$k\,\mathrm{P}(X_3 = k)$	1/3	1	1/2	$11/6 = E(X_3)$
$k^2\,\mathrm{P}(X_3 = k)$	1/3	2	3/2	$23/6 = E(X_3^2)$

d'où $V(X_3) = \dfrac{23}{6} - \dfrac{121}{36} = \dfrac{17}{36}$

3. a) " X_n prend ses valeurs dans $\{1, 2, ..., n\}$" peut être compris comme "ce sont toutes les valeurs de X_n " ou bien comme "X_n prend des valeurs parmi celles là" (plutôt cette deuxième forme)

Comme on retire au moins une boule à chaque tirage, on en fait au maximum n et au maximum, $X_n = n$

Et au minimum, on obtient la boule 1 au premier. Donc X_n prend ses valeurs dans $\{1, 2, ..., n\}$

b) Comme $(X_n = 1) = (I_n = 1)$ on a $\mathrm{P}(X_n = 1) = 1/n$

$(X_n = n)$ ne peut survenir que si l'on retire une seule boule à chaque tirage. Donc si pour tout k, on tire au $k^{ième}$ tirage la plus grande des boules

restantes. Donc $(X_n = k) = n_1 \cap (n-1)_2 \cap \cdots \cap 1_n$ en codant i_j l'événement "obtenir i au $j^{ème}$ tirage"

On a donc $\mathrm{P}(X_n = k) = \mathrm{P}(n_1)\,\mathrm{P}_{n_1}(n - 1_2)\ldots\mathrm{P}_{n_1\cap\cdots\cap 2_{n-1}}(1_n)$

le conditionnement donne les boules restantes dans l'urne (qui sont équiprobables) :

$\mathrm{P}(X_n = n) = \mathrm{P}(n_1)\,\mathrm{P}_{n-1\text{ boules}}(n - 1_2)\ldots\mathrm{P}_{1\text{ boule}}(1_n) = \dfrac{1}{n}\cdot\dfrac{1}{n-1}\cdots\dfrac{1}{1} = \dfrac{1}{n!}$

c) On connaît déjà $\mathrm{P}(X_n = 1)$ donc on ne s'intéresse à $\mathrm{P}(X_n = j)$ que pour $j \geq 2$.

Pour avoir $X_n = j$, on réutilise $\mathrm{P}_{I_n=k}(X_n = j) = \mathrm{P}(X_{k-1} = j - 1)$ via la formule des probabilités totales :

$(I_n = k)_{k\in[[1,n]]}$ est un système complet d'événements. Donc

$$\begin{aligned}
\mathrm{P}(X_n = j) &= \sum_{k=1}^{n}\mathrm{P}_{I_n=k}(X_n = j)\,\mathrm{P}(I_n = k)\\
&= \sum_{k=2}^{n}\mathrm{P}_{I_n=k}(X_n = j)\,\mathrm{P}(I_n = k) + \mathrm{P}_{I_n=1}(X_n = j)\,\mathrm{P}(I_n = 1)
\end{aligned}$$

Si $I_n = 1$ alors $X_n = 1$, et comme $j \geq 2$ on a $\mathrm{P}(X_n = j/I_n = 1) = 0$.

N.B. si $k < j$ et que $I_n = k$ il reste $k-1$ boules après le premier tirage donc on les aura épuisées en $k-1$ tirages au plus tard (donc au $k^{ième}$). On ne peut donc pas avoir $X_n = j$ et $\mathrm{P}(X_n = j/I_n = k) = 0$.

On aurait donc pu extraire de la somme tous les termes de 1 à $j-1$.

$$\begin{aligned}
\mathrm{P}(X_n = j) &= \sum_{k=2}^{n}\frac{1}{n}\mathrm{P}(X_{k-1} = j - 1)\\
\text{réindexé } h = k - 1 : \quad &= \frac{1}{n}\sum_{k=1}^{n-1}\mathrm{P}(X_k = j - 1)
\end{aligned}$$

Donc

$$\forall j \geqslant 2, \quad \mathrm{P}(X_n = j) = \frac{1}{n}\sum_{k=1}^{n-1}\mathrm{P}(X_k = j - 1)$$

d) Si n est supérieur ou égal à 3 (donc $n-1$ est supérieur à 2 et on peut réutiliser le précédent) et j supérieur ou égal à 2, on a :

$$\begin{aligned}
n\,\mathrm{P}(X_n = j) - (n-1)\,\mathrm{P}(X_{n-1} = j) &= \frac{n}{n}\sum_{k=1}^{n-1}\mathrm{P}(X_k = j - 1) - \sum_{k=1}^{n-2}\mathrm{P}(X_k = j - 1)\\
&= \sum_{k=1}^{n-1}\mathrm{P}(X_k = j - 1) - \sum_{k=1}^{n-2}\mathrm{P}(X_k = j - 1)\\
&= \mathrm{P}(X_{n-1} = j - 1)
\end{aligned}$$

Donc pour $n \geq 3$ et $j \geq 2$ on a

$$\begin{aligned} n\,\mathrm{P}(X_n = j) &= (n-1)\,\mathrm{P}(X_{n-1} = j) + \mathrm{P}(X_{n-1} = j-1) \text{ donc} \\ \mathrm{P}(X_n = j) &= \frac{n-1}{n}\mathrm{P}(X_{n-1} = j) + \frac{1}{n}\mathrm{P}(X_{n-1} = j-1) \end{aligned}$$

Pour $j = 1$ on a pour tout entier $n \geq 1$: $\mathrm{P}(X_n = 1) = \dfrac{1}{n}$ donc

$$\frac{n-1}{n}\mathrm{P}(X_{n-1} = 1) + \frac{1}{n}\mathrm{P}(X_{n-1} = 0) = \frac{n-1}{n}\frac{1}{n-1} + \frac{1}{n}0 = \frac{1}{n} = \mathrm{P}(X_n = 1)$$

et la propriété est encore vraie pour $j = 1$
Enfin pour $n = 2$, le seul cas non prouvé est $j = 2$ ($j = 1$ est déjà traité et $j \geq 3$ donne des probabilité nulles)

$$\frac{2-1}{2}\mathrm{P}(X_1 = 2) + \frac{1}{2}\mathrm{P}(X_1 = 1) = 0 + \frac{1}{2} = \mathrm{P}(X_2 = 2)$$

Donc si $n \geq 2$

$$\forall j \geqslant 1\ ,\ \mathrm{P}(X_n = j) = \frac{n-1}{n}\mathrm{P}(X_{n-1} = j) + \frac{1}{n}\mathrm{P}(X_{n-1} = j-1)$$

4. **a)** Comme $X_n(\Omega) \subset [[1, n]]$, on a

$$\begin{aligned} E(X_n) &= \sum_{k=1}^{n} j\,\mathrm{P}(X_n = j) \\ &= \sum_{j=1}^{n} j\left[\frac{n-1}{n}\mathrm{P}(X_{n-1} = j) + \frac{1}{n}\mathrm{P}(X_{n-1} = j-1)\right] \\ &= \frac{n-1}{n}\sum_{j=1}^{n} j\,\mathrm{P}(X_{n-1} = j) + \frac{1}{n}\sum_{j=1}^{n} j\,\mathrm{P}(X_{n-1} = j-1) \end{aligned}$$

On enlève de la première somme le terme pour $j = n$: $\mathrm{P}(X_{n-1} = n) = 0$ et on réindexe la seconde par $k = j - 1$

$$\begin{aligned} E(X_n) &= \frac{n-1}{n}\sum_{j=1}^{n-1} j\,\mathrm{P}(X_{n-1} = j) + 0 + \frac{1}{n}\sum_{k=0}^{n-1}(k+1)\,\mathrm{P}(X_{n-1} = k) \\ &= \frac{n-1}{n}E(X_{n-1}) + \frac{1}{n}\sum_{k=1}^{n-1} k\,\mathrm{P}(X_{n-1} = k) + \frac{1}{n}\sum_{k=1}^{n-1}\mathrm{P}(X_{n-1} = k) + 0 \\ &= \frac{n-1}{n}E(X_{n-1}) + \frac{1}{n}E(X_{n-1}) + \frac{1}{n} \\ &= E(X_{n-1}) + \frac{1}{n} \end{aligned}$$

b) Comme $E(X_1) = 1$, on a $E(X_n) = 1 + \dfrac{1}{2} + \cdots + \dfrac{1}{n} = \displaystyle\sum_{k=1}^{n}\frac{1}{k}$

Et $E(X_n) = h_n \sim \ln(n)$ quand n tend vers l'infini.

5. a) On réutilise la méthode précédente :

$$\begin{aligned}
E\left(X_n^2\right) &= \sum_{k=1}^{n} j^2\,\mathrm{P}\left(X_n = j\right)\\
&= \sum_{j=1}^{n} j^2\left[\frac{n-1}{n}\,\mathrm{P}\left(X_{n-1}=j\right)+\frac{1}{n}\,\mathrm{P}\left(X_{n-1}=j-1\right)\right]\\
&= \frac{n-1}{n}\sum_{j=1}^{n-1} j^2\,\mathrm{P}\left(X_{n-1}=j\right)+0+\frac{1}{n}\sum_{j=2}^{n} j^2\,\mathrm{P}\left(X_{n-1}=j-1\right)+0\\
&= \frac{n-1}{n}E\left(X_{n-1}^2\right)+\frac{1}{n}\sum_{k=1}^{n-1}\left(k^2+2k+1\right)\mathrm{P}\left(X_{n-1}=k\right)\\
&= \frac{n-1}{n}E\left(X_{n-1}^2\right)+\frac{1}{n}\sum_{k=1}^{n-1} k^2\,\mathrm{P}\left(X_{n-1}=k\right)\\
&\quad+\frac{2}{n}\sum_{k=1}^{n-1} k\,\mathrm{P}\left(X_{n-1}=k\right)+\frac{1}{n}\sum_{k=1}^{n-1}\mathrm{P}\left(X_{n-1}=k\right)\\
&= \frac{n-1}{n}E\left(X_{n-1}\right)+\frac{1}{n}E\left(X_{n-1}^2\right)+\frac{2}{n}E\left(X_{n-1}\right)+\frac{1}{n}\\
&= E\left(X_{n-1}^2\right)+\frac{2}{n}E\left(X_{n-1}\right)+\frac{1}{n}
\end{aligned}$$

b) Comme $V\left(X_n\right)=E\left(X_n^2\right)-E\left(X_n\right)^2$ on a

$$\begin{aligned}
V\left(X_n\right) &= E\left(X_{n-1}^2\right)+\frac{2}{n}E\left(X_{n-1}\right)+\frac{1}{n}-\left(E\left(X_{n-1}\right)+\frac{1}{n}\right)^2\\
&= E\left(X_{n-1}^2\right)+\frac{2}{n}E\left(X_{n-1}\right)+\frac{1}{n}-\left(E\left(X_{n-1}\right)^2+\frac{2}{n}E\left(X_{n-1}\right)+\frac{1}{n^2}\right)\\
&= V\left(X_{n-1}\right)+\frac{1}{n}-\frac{1}{n^2}
\end{aligned}$$

Et comme $V\left(X_1\right)=0$ on a alors

$$\begin{aligned}
V(X_n) &= 0+\frac{1}{2}-\frac{1}{4}+\frac{1}{3}-\frac{1}{9}+\cdots+\frac{1}{n}-\frac{1}{n^2}\\
&= \sum_{k=2}^{n}\frac{1}{k}-\sum_{k=2}^{n}\frac{1}{k^2}\\
&= \sum_{k=1}^{n}\frac{1}{k}-\sum_{k=1}^{n}\frac{1}{k^2}\\
&= h_n - k_n
\end{aligned}$$

c) Donc $V(X_n)\sim \ln(n)$ quand n tend vers l'infini.

6. Soit $\left(T_i\right)_{i\geqslant 1}$une suite de variables aléatoires indépendantes telle que, pour tout i entier naturel non nul, T_i suit la loi de Bernoulli de paramètre $\dfrac{1}{i}$. On pose :

$$S_n=\sum_{i=1}^{n} T_i = T_1+.....+T_n$$

a) T_1 suit une loi de Bernoulli de paramètre 1 donc $\mathrm{P}(T_1 = 1) = 1 = \mathrm{P}(X_1 = 1)$ et $\mathrm{P}(T_1 = 0) = 0 = \mathrm{P}(X_1 = 0)$

Donc X_1 et T_1 ont même loi.

b) On a pour tout entier j (même si $j = 0$ auquel cas tous les événements sont impossibles) $(S_n = j) = \left(\sum\limits_{i=1}^{n} T_i = j\right)$ et comme les valeurs possibles de T_n ne sont que 0 et 1,

$$\begin{aligned}\left(\sum_{i=1}^{n} T_i = j\right) &= \left[\left(\sum_{i=1}^{n-1} T_i = j\right) \cap (T_n = 0)\right] \cup \left[\left(\sum_{i=1}^{n-1} T_i = j-1\right) \cap (T_n = 1)\right] \\ &= [(S_{n-1} = j) \cap (T_n = 0)] \cup [(S_{n-1} = j-1) \cap (T_n = 1)]\end{aligned}$$

Les deux étant incompatibles et S_{n-1} (qui ne dépend que de $T_1 \ldots, T_{n-1}$) étant indépendant de T_n

$$\begin{aligned}\mathrm{P}(S_n = j) &= \mathrm{P}[(S_{n-1} = j) \cap (T_n = 0)] + \mathrm{P}[(S_{n-1} = j-1) \cap (T_n = 1)] \\ &= \mathrm{P}(S_{n-1} = j-1)\,\mathrm{P}(T_n = 0) + \mathrm{P}(S_{n-1} = j-1)\,\mathrm{P}(T_n = 1) \\ &= \frac{n-1}{n}\,\mathrm{P}(S_{n-1} = j) + \frac{1}{n}\,\mathrm{P}(S_{n-1} = j-1)\end{aligned}$$

C.Q.F.D

On a alors par récurrence sur n que pour tout entier j, (et on devra traiter à part le cas $j = 1$ qui fait intervenir $S_{n-1} = 0$ si on n'a pas inclus la valeur 0 au dessus) que $\mathrm{P}(S_n = j) = \mathrm{P}(X_n = j)$ donc que X_n et S_n ont même loi.

c) Comme X_n et S_n ont la même loi, elles ont la même variance et la même espérance. Or

$$E(S_n) = \sum_{i=1}^{n} E(T_i) = \sum_{i=1}^{n} \frac{1}{i} = h_n$$

et comme les (T_i) sont indépendantes,

$$\begin{aligned}V(S_n) &= \sum_{i=1}^{n} V(T_i) = \sum_{i=1}^{n} \frac{1}{i}\left(1 - \frac{1}{i}\right) = \sum_{i=1}^{n} \frac{1}{i} - \sum_{i=1}^{n} \frac{1}{i^2} \\ &= h_n - k_n\end{aligned}$$

Et on retrouve donc bien $E(X_n) = h_n$ et $V(X_n) = h_n - k_n$

Partie 3 : Etude de la variable aléatoire $\mathbf{Y}_n$.

1. Pour Y_1, comme la seule boule dans l'urne est 0, Y_1 et la variable certaine égale à 1.

2. **a)** Pour Y_2 on a dans l'urne les boules 1 et 2. Les sommes possibles sont donc : 1 (1 tiré en premier) et 2+1 (2 tiré en premier puis 1)

Donc $Y_2(\Omega) = \{1, 3\}$

b) Comme $(Y_2 = 1) = 1_1$ (on tire 1 en premier) on a $\mathrm{P}(Y_1 = 1) = 1/2$

et $\mathrm{P}(Y_2 = 3) = \mathrm{P}(2_1 \cap 1_2) = \mathrm{P}(2_1)\,\mathrm{P}(1_1/2_1) = 1/2 \cdot 1 = 1/2$

a) Si on tire la boule k en premier, il reste les boules jusqu'à $k-1$.

Donc pour faire un total de j, il reste à faire un total de $j-k$ avec les $k-1$ boules restantes.

Donc si n est supérieur ou égal à 2, pour tout entier j non nul et tout entier k supérieur ou égal à 2

$$\mathrm{P}_{I_n=k}\left(Y_n=j\right)=\mathrm{P}\left(Y_{k-1}=j-k\right)$$

On a besoin de $k\geq 2$ pour que la variable Y_{k-1} soit bien définie.

(si $j<k$, on ne pourra pas obtenir un total plus petit que le premier tirage, ce que l'on retrouve bien dans cette formule où toutes les probabilités seront nulles)

b) On a alors par la formule des probabilités totales :

$(I_n=k)_{k\in[[1,n]]}$ est un système complet d'événements donc

$$\begin{aligned}\mathrm{P}\left(Y_n=j\right) &= \sum_{k=1}^{n}\mathrm{P}_{I_n=k}\left(Y_n=j\right)\mathrm{P}\left(I_n=k\right)\\ &= \sum_{k=2}^{n}\mathrm{P}_{I_n=k}\left(Y_n=j\right)\mathrm{P}\left(I_n=k\right)+\mathrm{P}_{I_n=1}\left(Y_n=j\right)\mathrm{P}\left(I_n=1\right)\end{aligned}$$

on doit traiter à part le cas où $k=1$ et la probabilité sera nulle pour $j>1$

- Pour $j=1$ on a $\mathrm{P}\left(Y_n=1\right)=\dfrac{1}{n}$ (il faut obtenir 1 dès le premier tirage)
  $\dfrac{n-1}{n}\mathrm{P}\left(Y_{n-1}=1\right)+\dfrac{1}{n}\mathrm{P}\left(Y_{n-1}=1-n\right)=\dfrac{n-1}{n}\dfrac{1}{n-1}+0=\dfrac{1}{n}$ donc la formule est vraie pour $j=1$
- On traite à présent uniquement pour $j>1$: on a

$$\mathrm{P}\left(Y_n=j\right)=\sum_{k=2}^{n}\mathrm{P}\left(Y_{k-1}=j-k\right)\frac{1}{n}$$

  Comme n n'intervient plus que comme borne supérieure de la somme et dans le $1/n$, on a donc

$$\mathrm{P}\left(Y_{n-1}=j\right)=\sum_{k=2}^{n-1}\mathrm{P}\left(Y_{k-1}=j-k\right)\frac{1}{n-1}$$

  Donc

$$\begin{aligned}\frac{n-1}{n}\mathrm{P}\left(Y_{n-1}=j\right)+\frac{1}{n}\mathrm{P}\left(Y_{n-1}=j-n\right) &= \frac{1}{n}\sum_{k=2}^{n-1}\mathrm{P}\left(Y_{k-1}=j-k\right)+\frac{1}{n}\mathrm{P}\left(Y_{n-1}=j-n\right)\\ &= \frac{1}{n}\sum_{k=2}^{n}\mathrm{P}\left(Y_{k-1}=j-k\right)\frac{1}{n}\\ &= \mathrm{P}\left(Y_n=j\right)\end{aligned}$$

c) On calcule l'espérance de Y_n en réutilisant la relation précédente :

$$\begin{aligned}E\left(Y_n\right) &= \sum_{j\in Y_n(\Omega)} j\,\mathrm{P}\left(Y_n=j\right)\\ &= \frac{n-1}{n}\sum_{j\in Y_n(\Omega)} j\,\mathrm{P}\left(Y_{n-1}=j\right)+\frac{1}{n}\sum_{j\in Y_n(\Omega)} j\,\mathrm{P}\left(Y_{n-1}=j-n\right)\end{aligned}$$

Les valeur possibles de Y_n sont d'une part les mêmes que celles de Y_{n-1}(si on ne commence pas par n)
et d'autre part celles Y_{n-1} auguementées de n (si on tire n en premier)
Donc quand j parcours $Y_n(\Omega)$, il prend toutes les valeurs de $Y_{n-1}(\Omega)$ et $j-n$ prendra également toutes les valeurs de $Y_{n-1}(\Omega)$
(si $j \notin Y_{n-1}(\Omega)$ alors $\mathrm{P}(Y_{n-1}=j)=0$ et si $j-n \notin Y_{n-1}(\Omega)$ alors $\mathrm{P}(Y_{n-1}=j-n)=0$)
Finalement en réindexant la deuxième somme par $i=j-k$

$$\begin{aligned} E(Y_n) &= \frac{n-1}{n} \sum_{j \in Y_{n-1}(\Omega)} j\,\mathrm{P}(Y_{n-1}=j)+0+\frac{1}{n} \sum_{i \in Y_{n-1}(\Omega)} (i+n)\,\mathrm{P}(Y_{n-1}=i)+0 \\ &= \frac{n-1}{n} E(Y_{n-1})+\frac{1}{n} E(Y_{n-1})+\frac{n}{n} \sum_{i \in Y_{n-1}(\Omega)} \mathrm{P}(Y_{n-1}=i) \\ &= E(Y_{n-1})+1 \end{aligned}$$

Comme $E(Y_1)=1$ et que l'on a une suite arithmétique de raison 1 on a $E(Y_n)=(n-1)+E(Y_1)=n$

Partie 4. Simulation informatique.

Dans la procédure Truc, récursive, a et b sont des accumulateurs qui calculent X_n et Y_n (raison pour laquelle elles sont déclarées var) et alea représente la boule tirée au hasard.
En effet :

- on tire un nombre (boule) au hasard entre 1 et n .... alea : = random (n) + 1
- on affiche le numéro obtenu .... writeln (alea)
- on compte un tirage de plus .... a : = a+l
- on accumule le numéro .... b : = b + alea
- puis on recommence avec les boules restantes (numéros $\leq alea-1$) .... Truc (alea- l , a , b)
- Il s'arrète quand le numéro 1 est obtenu (et ne totalise pas ce numéro)

le programme

- demande le nombre total de boules
- initialise les compteurs à 1 (et pas à 0...?) car le tirage avec la boule 1 n'est pas comptabilisé (les compteurs étant après la condition .... if alea > 1 then.
- fait les tirages .... Truc (n , a , b) ;
- puis affiche les résultats .... writeln ('a = ‘, a , 'b = ', b)

Avec un programme itératif :

```
var n , a , b :  integer ;
Begin
a :  =0 ; b :  = 0; .
write (‘ n :  ’)  ; readln (n);
repeat
    n :  = random (n) + 1 ;
    writeln (n) ;
    a :  = a+1;
    b :  = b + n;
until n=1
writeln (’a = ‘, a , ’b = ’, b),
End.
```

ce qui est beaucoup plus simple et naturel.

Partie 5.

On considère l'urne U_n contenant n boules numérotées entre 1 et n. A partir de l'urne U_n on effectue la suite de tirages décrite dans l'entête du problème. Pour i entier de $\{1, ..., n\}$, on définit $Z_i^{(n)}$ la variable aléatoire égal à 1 si, lors d'un quelconque de ces tirages, on a obtenu la boule numéro i, égale à 0 sinon.

1. La boule n ne peut être obtenue que lors du premier tirage.

Donc$\left(Z_n^{(n)} = 1\right) = n_1$ et $\mathrm{P}\left(Z_n^{(n)} = 1\right) = \dfrac{1}{n}$ et $\mathrm{P}\left(Z_n^{(n)} = 0\right) = \dfrac{n-1}{n}$ donc $Z_n^{(n)}$ suit une loi de Bernoulli de paramètre $1/n$

La variable $Z_1^{(n)}$ vaut 1 si le numéro 1 est obtenu lors des tirages. Or l'expérience s'arrête justement au tirage de 1. Donc $Z_1^{(n)}$ est la variable constante qui vaut 1.

2. **a)** Comme précédemment on distingue suivant la valeur de I_n :

- Si $I_n = k$ et que $k < i$, on retire toutes les boules$\geq k$ - donc la boule i - et on ne peut plus l'obtenir
  Donc $\mathrm{P}\left(Z_i^{(n)} = 1/I_n = k\right) = 0$
- Si $I_n = j$ on a eu j dès le premier tirage donc $\mathrm{P}\left(Z_i^{(n)} = 1/I_n = j\right) = 1$
- Si $I_n = k$ et que $k \geq i$, on continue l'expérience avec les boules $\leq k-1$.
  Donc $\mathrm{P}\left(Z_i^{(n)} = 1/I_n = k\right) = \mathrm{P}\left(Z_i^{(k-1)} = 1\right)$

En utilisant la formule des probabilités totales, on trouve alors : $(I_n = k)_{k\in[[1,n]]}$ système complet d'événement et i un entier de $\{1, ..., n-1\}$

$$\begin{aligned}
\mathrm{P}\left(Z_i^{(n)} = 1\right) &= \sum_{k=1}^{n} \mathrm{P}_{I_n=k}\left(Z_i^{(n)} = 1\right) \mathrm{P}\left(I_n = k\right) \\
&= \sum_{k=1}^{i-1} 0 + \mathrm{P}_{I_n=i}\left(Z_i^{(n)} = 1/I_n = i\right) \mathrm{P}\left(I_n = i\right) + \sum_{k=i+1}^{n} \mathrm{P}_{I_n=k}\left(Z_i^{(n)} = 1/I_n = k\right) \mathrm{P}\left(\right. \\
&= \frac{1}{n} + \sum_{k=i+1}^{n} \frac{1}{n} \mathrm{P}\left(Z_i^{(k-1)} = 1\right)
\end{aligned}$$

b) Pour calculer $\mathrm{P}\left(Z_i^{(n)}=1\right)$ on a besoin de $\mathrm{P}\left(Z_i^{(k-1)}=1\right)$ pour tous les k de $i+1$ à n.

On prend donc comme hypothèse de récurrence et on démonter que pour tout $n\geq 1$, on a pour tout $1\leq k\leq n$ et pour tout i de $\{1,...,k\}$, $Z_i^{(k)}$ suit la loi de Bernoulli de paramètre $\dfrac{1}{i}$.:

C'est vrai pour $n=1$ et $i\in\{1\}$, $Z_i^{(1)}=Z_1^{(1)}$ est certaine égale à 1 donc suit une loi de Bernoulli de paramètre $1=1/1$

Soit $n\geq 1$ tel que pour tout $n\geq 1$, on a et pour tout i de $\{1,...,n\}$ et pour tout $1\leq k\leq n$, $Z_i^{(k)}$ suit la loi de Bernoulli de paramètre $\dfrac{1}{i}$.

Alors pour $n+1$ il y a à démontrer la propriété pour $k=n+1$ (les autres cas sont dans l'hypothèse) :

on peut utiliser la propriété du 2a pour $i\in\{1\ldots,n\}$, la valeur $i=n+1$ devra être traitée à part.

- Pour $i\in\{1\ldots,n\}$ on a $\mathrm{P}\left(Z_i^{(n+1)}=1\right)=\dfrac{1}{n+1}+\sum\limits_{k=i+1}^{n+1}\dfrac{1}{n+1}\mathrm{P}\left(Z_i^{(k-1)}=1\right)$

  et comme $i+1\leq k\leq n+1$ alors $i\leq k-1\leq n$ et $Z_i^{(k-1)}$ par hypothèse de récurrence suit une loi de Bernoulli de paramètre $\dfrac{1}{i}$. Donc

$$\begin{aligned}
\mathrm{P}\left(Z_i^{(n+1)}=1\right) &= \frac{1}{n+1}+\sum_{k=i+1}^{n+1}\frac{1}{n+1}\mathrm{P}\left(Z_i^{(k-1)}=1\right)\\
&= \frac{1}{n+1}\left[1+\sum_{k=i+1}^{n+1}\frac{1}{i}\right]\\
&= \frac{1}{n+1}\frac{i+(n+1-i-1+1)}{i}\\
&= \frac{1}{i}
\end{aligned}$$

  Donc pour tout $i\in\{1\ldots,n\}$ la variable aléatoire $Z_i^{(n+1)}$ suit bien une loi de Bernoulli de paramètre $\dfrac{1}{i}$

- et pour $i=n+1$ on a d'après 1 $Z_{n+1}^{(n+1)}$ qui suit une loi de Bernoulli de paramètre $\dfrac{1}{n+1}$

C.Q.F.D.

Donc pour tout entier n, on a pour tout $i\in\{1,\ldots,n\}$ $Z_i^{(n)}$ qui suit une loi de Bernoulli de paramètre $\dfrac{1}{i}$

3. $Z_i^{(n)}$ est le nombre de boules i obtenue lors des tirages (on en a eu 1 ou 0)

Donc $\sum\limits_{i=1}^{n}Z_i^{(n)}$ est le nombre total de boules obtenues, tous numéros confondus.

Donc $\sum\limits_{i=1}^{n}Z_i^{(n)}=X_n$ et $E(X_n)=\sum\limits_{i=1}^{n}E\left(Z_i^{(n)}\right)=\displaystyle\sum_{i=1}^{n}\frac{1}{i}=h_n$ car l'espérance d'une variable suivant une loi de Bernoulli de paramètre $\dfrac{1}{i}$ est $\dfrac{1}{i}$.

Et on retrouve bien le résultat vu précédemment.

4. On a $Y_n = \sum_{i=1}^{n} i \cdot Z_i^{(n)}$

En effet si la boule i est tirée, alors $Z_i^{(n)} = 1$ et on rajoute $i \cdot 1 = i$ et on ne rajoute rien au total sinon.

Donc
$$E(Y_n) = \sum_{i=1}^{n} E\left(i \cdot Z_i^{(n)}\right) = \sum_{i=1}^{n} i \cdot E\left(Z_i^{(n)}\right) = \sum_{i=1}^{n} \frac{i}{i} = \sum_{i=1}^{n} 1 = n$$